

## BUNDEREPUBLIK DEUTSCHLAND

09/462863



REC'D	22 OCT 1998
WIPO	PCT

PRIORITY  
DOCUMENTSUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

## Bescheinigung

Die Robert Bosch GmbH in Stuttgart/Deutschland hat eine  
Patentanmeldung unter der Bezeichnung

"Verfahren zur Erzeugung eines verbesserten Bildsignals bei der Bewegungsschätzung von Bildsequenzen, insbesondere eines Prädiktionssignals für Bewegungsbilder mit bewegungskompensierender Prädiktion"

am 15. Juli 1997 beim Deutschen Patentamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patentamt vorläufig das Symbol H 04 N 5/14 der Internationalen Patentklassifikation erhalten.

München, den 3. August 1998

Der Präsident des Deutschen Patentamts

Im Auftrag

Aktenzeichen: 197 30 305.6

11.07.97 Sk/Os

5

ROBERT BOSCH GMBH, 70442 Stuttgart

10

Verfahren zur Erzeugung eines verbesserten Bildsignals bei  
der Bewegungsschätzung von Bildsequenzen, insbesondere eines  
Prädiktionssignals für Bewegtbilder mit  
bewegungskompensierender Prädiktion

15

Stand der Technik

20

Die Erfindung geht aus von einem Verfahren zur Erzeugung  
eines verbesserten Bildsignals bei der Bewegungsschätzung  
von Bildsequenzen, insbesondere eines Prädiktionssignals für  
Bewegtbilder mit bewegungskompensierender Prädiktion, wobei  
für Bildblöcke Bewegungsvektoren gebildet werden, die für  
jeden Bildblock eines aktuellen Bildes die Position des zur  
Prädiktion benutzten Bildblockes gegenüber einem zeitlich  
davorliegenden Referenzbild angeben.

25

30

Aus der EP 0 558 922 B2 ist ein Verfahren zur Verbesserung  
der Bewegungsschätzung in Bildsequenzen in  
Halbpelgenauigkeit nach dem Full-Search-Verfahren bekannt.  
Dort wird in einem ersten Verfahrensschritt der Suchbereich  
und in einem zweiten Verfahrensschritt der Match-Block  
gefiltert unter Zuhilfenahme eines zusätzlichen digitalen  
Filters, daß eine Rasterverschiebung des Bildpunkttrasters um  
 $\frac{1}{4}$  pel ermöglicht. Mit dieser Maßnahme ist eine Verfälschung  
des Bewegungsvektorfeldes auszuschließen.

35

Beim „MPEG-4 Video Verification Model Version 7.0“, Bristol, April 1997, MPEG 97/N1642 in ISO/IEC JTC1/SC 29/WG11 ist ein Encoder und Decoder zur objektbasierten Codierung von Bewegtbildfolgen spezifiziert. Dabei werden nicht mehr  
5 rechteckige Bilder fester Größe codiert und zum Empfänger übertragen, sondern sogenannte VIDEO OBJECTS (VO), welche beliebige Form und Größe aufweisen dürfen. Die Abbildung eines solchen VO in der Kamerabildebene zu einem bestimmten Zeitpunkt wird als VIDEO OBJECTS PLANE (VOP) bezeichnet. Somit ist die Beziehung zwischen VO und VOP äquivalent zu  
10 der Beziehung zwischen Bildfolge und Bild im Falle der Übertragung rechteckiger Bilder fester Größe.

Die bewegungskompensierende Prädiktion im Verifikationsmodell wird mit Hilfe sogenannter blockweiser Bewegungsvektoren durchgeführt, die für jeden Block der Größe 8 x 8 bzw. 16 x 16 Bildpunkte des aktuellen Bildes die Position des zur Prädiktion benutzten Blockes in einem bereits übertragenen Referenzbild angeben. Die Auflösung der  
15 Bewegungsvektoren ist dabei auf einen halben Bildpunkt beschränkt, wobei Bildpunkte zwischen dem Abtastraster (half pixel position) durch eine bilineare Interpolationsfilterung aus den Bildpunkten auf dem Abtastraster (integer pixel position) erzeugt werden (Figur 1).+ gibt hierbei die  
20 Ganzzahl-Pixel-Position und 0 die Halb-Pixel-Position an. Die interpolierten Werte a, b, c und d in Halb-Pixel-Position ergeben sich durch folgende Beziehungen:  
$$a = A, b = (A + B) // 2, c = (A + C) // 2,$$
$$d = (A + B + C + D) // 4, \text{ wobei } // \text{ eine gerundete Ganzzahl-Division kennzeichnet.}$$
  
30

#### Vorteile der Erfindung

Mit den Maßnahmen der Erfindung läßt sich die Qualität des  
35 Prädiktionssignals und somit die Kodierungseffizienz

verbessern. Dabei wird zur Erzeugung von Bildpunkten zwischen dem Bildpunkt-Abtastraster eine größere örtliche Nachbarschaft berücksichtigt, als bei der bilinearen Interpolation. Die erfindungsgemäße aliasing-reduzierende Interpolationsfilterung führt zu einer erhöhten Auflösung des Bewegungsvektors und damit zu einem Prädiktionsgewinn und einer erhöhten Codierungseffizienz. Bei der Erfindung können die FIR-Filterkoeffizienten den zu codierenden Signalen angepaßt und für jedes Video-Object getrennt übertragen werden, was eine weitere Effizienzerhöhung für die Codierung ermöglicht sowie die Flexibilität des Verfahrens erhöht.

Im Gegensatz zur Lösung gemäß der EP 0 558 922 B1 müssen keine zusätzlichen Polyphasenfilterstrukturen für Zwischenpositionen mit  $\frac{1}{4}$  pel-Bildpunktauflösung in horizontaler und vertikaler Richtung entworfen werden.

Mit den Maßnahmen der Erfindung ist es möglich, daß bei gleichbleibender Datenrate die Bildfolgefrequenz eines MPEG-1-Coders von 25 Hz auf 50 Hz verdoppelt werden kann. Bei einem MPEG2-Coder kann die Datenrate bei gleichbleibender Bildqualität um bis zu 30% reduziert werden.

## Zeichnungen

Anhand von Zeichnungen werden nun Ausführungsbeispiele der Erfindung näher erläutert. Es zeigen:

Figur 2 ein Interpolationsschema für Bildpunkte zwischen dem Bildpunkt-Abtastraster,  
Figur 3 die Struktur eines FIR-Filters zur Interpolation,  
Figur 4 eine weitere Interpolation mit noch höherer Auflösung.

## Beschreibung von Ausführungsbeispielen

Beim Verfahren nach der Erfindung werden für Bildblöcke Bewegungsvektoren gebildet, die für jeden Bildblock eines  
5 aktuellen Bildes die Position des zur Prädiktion benutzten Bildblockes gegenüber einem zeitlich davorliegenden Referenzbild angeben.

Die Bestimmung der Bewegungsvektoren für die Prädiktion erfolgt in drei aufeinanderfolgenden Schritten:

10 In einem ersten Suchschritt wird ein Bewegungsvektor mit Pelgenauigkeit nach einem herkömmlichen Verfahren, z.B. nach der Full-Search-Block-Matching-Methode, für jeden Bildblock bestimmt. Hierbei wird das Minimum des Fehlerkriteriums für  
15 mögliche Bewegungspositionen ermittelt und jener Vektor, der die Bewegung des Bildblockes am besten beschreibt, ausgewählt (EP 0 368 151 B1).

20 In einem zweiten Suchschritt, der wiederum auf einer solchen Minimumsuche für das Fehlerkriterium basiert, wird ausgehend von dem im ersten Schritt ermittelten Bewegungsvektor mit Pelgenauigkeit durch eine aliasing-reduzierende Interpolationsfilterung mittels eines digitalen  
25 symmetrischen FIR (finite impulse response)-Filters ein verbesserter Bewegungsvektor auf Subpelgenauigkeit ermittelt. Die Auflösung wird hierbei höher gewählt als im ersten Suchschritt, vorzugsweise wird eine Auflösung von einem halben Bildpunkt bezogen auf das Bildpunktraster gewählt. Figur 2 zeigt das Interpolationsmuster für die  
30 Bildpunkte b, c und d zwischen dem Bildpunktraster, die sich aus den Nachbarbildpunkten A, B, C, D, E, F, G, H auf dem Bildpunktraster ergeben. + gibt die Ganzzahl-Pixelposition an, o die Halbpixelposition. Es gilt:

$$b = (CO1x(A_{-1} + A_{+1}) + CO2x(A_{-2} + A_{+2}) + CO3x(A_{-3} + A_{+3}) + CO4x(A_{-4} + A_{+4}))/256$$

$$c_i = (CO1x(A_i + E_i) + CO2x(B_i + F_i) + CO3x(C_i + G_i) + CO4x(D_i + H_i))/256$$

$$5 \quad d = (CO1x(c_{-1} + c_{+1}) + CO2x(c_{-2} + c_{+2}) + CO3x(c_{-3} + c_{+3}) + CO4x(c_{-4} + c_{+4}))/256$$

Die Struktur des verwendeten FIR-Interpolationsfilters ist aus Figur 3 ersichtlich. Es weist nach jeder Bildpunktposition  $\delta p$  einen Abzweig zu einem Koeffizientenbewerter 1, 2, 3 usw. auf und eine Summationseinrichtung 10 am Ausgang. Wie aus obigen Beziehungen ersichtlich ist, wird zur Erzeugung von Bildpunkten zwischen dem Abtastraster eine größere örtliche Nachbarschaft berücksichtigt, als bei der bilinearen Interpolation nach dem Stand der Technik. Die Interpolationsfilterkoeffizienten CO2, CO3, CO4 werden dabei so bestimmt, daß die Interpolationsfehlerleistung minimal wird. Die Koeffizienten können direkt mit dem bekannten Schätzverfahren des kleinsten, mittleren quadratischen Fehlers bestimmt werden. Aus der Minimierung der Interpolationsfehlerleistung erhält man ein lineares Gleichungssystem, dessen Koeffizienten aus dem Orthogonalitätsprinzip abgeleitet werden können. Ein derart optimierter Satz von FIR-Filterkoeffizienten ist durch die Koeffizienten CO1 = 161/256, CO2 = -43/256, CO3 = 23/256, CO4 = -8/256 gegeben.

Im dritten Suchschritt wird ausgehend von dem auf  $\frac{1}{2}$  Pelgenauigkeit bestimmten Bewegungsvektor durch eine weitere Interpolationsfilterung eine lokale Suche unter Zugrundelegung der acht Nachbarbildpunkte mit einer Auflösung, die noch weiter erhöht ist, vorzugsweise auf  $\frac{1}{4}$  Bildpunkt, durchgeführt. Es wird wie zuvor der Bewegungsvektor ausgewählt, der die geringste Prädiktionsfehlerleistung liefert.

Das Interpolationsmuster hierzu zeigt Figur 4. Die ganzzahligen Bildpunkt-Positionen sind mit X gekennzeichnet, die Halb-Pixel-Bildpunktpositionen mit o und die Viertel-Pixel-Bildpunktpositionen mit -. O markiert die beste Kompensation mit 1/2-Bildpunkt und + die Viertelbildpunkt-Suchposition.

Interpoliert wird bezüglich des Bildpunktrasters mit der Auflösung von einem halben Bildpunkt aus dem zweiten Suchschritt mit den FIR-Filterkoeffizienten  $CO1' = 1/2$ ,  $CO2' = 0$ ,  $CO3' = 0$ ,  $CO4' = 0$ .

Die gleiche zuvor vorgestellte Interpolationstechnik wird für die bewegungskompensierende Prädiktion verwendet.

Falls die Verarbeitung innerhalb eines Coders mit einem reduzierten Bildformat durchgeführt wird (SIF-Format innerhalb eines MPEG1-Coders oder Q-CIF in einem H.263-Coder), zur Anzeige aber das ursprüngliche Eingangsformat verwendet wird, z.B. CCIR 601[1] bei MPEG-1 oder CIF bei H.263, muß als Nachverarbeitung eine örtliche Interpolationsfilterung durchgeführt werden. Auch für diesen Zweck kann die beschriebene aliasing-kompensierende Interpolationsfilterung verwendet werden.

Um die aliasing-kompensierende Interpolation mit 1/4 Pel Auflösung zu aktivieren, können in einen Bildübertragungs-Bitstrom Aktivierungsbits eingefügt werden.

Für die Prädiktion von Video-Objekten können die Filterkoeffizienten  $CO1$  bis  $CO4$  und  $CO1'$  bis  $CO4'$  für jedes der Video-Objekte VO getrennt aufbereitet werden und in den Bildübertragungs-Bitstrom zu Beginn der Übertragung des jeweiligen Video-Objekts eingefügt werden.

Zur Codierung eines Bewegungsvektors kann der Wertebereich der zu codierenden Bewegungsvektor-Differenzen an die erhöhte Auflösung angepaßt werden.



11.07.97 Sk/Os

5

ROBERT BOSCH GMBH, 70442 Stuttgart

10

### Ansprüche

15

1. Verfahren zur Erzeugung eines verbesserten Bildsignals bei der Bewegungsschätzung von Bildsequenzen, insbesondere eines Prädiktionssignals für Bewegtbilder mit bewegungskompensierender Prädiktion, wobei für Bildblöcke Bewegungsvektoren gebildet werden, die für jeden Bildblock eines aktuellen Bildes die Position des zur Prädiktion benutzten Bildblockes gegenüber einem zeitlich davorliegenden Referenzbild angeben, mit folgenden

20

Schritten:

25

- In einem ersten Suchschritt wird ein Bewegungsvektor auf Pelgenauigkeit bestimmt,
- in einem zweiten Suchschritt wird ausgehend von dem Bewegungsvektor mit Pelgenauigkeit durch eine aliasing-reduzierende Interpolationsfilterung mittels eines digitalen Filters ein verbesserter Bewegungsvektor auf Subpelgenauigkeit ermittelt, wobei die Auflösung höher gewählt ist, als es der Auflösung des Bildpunktrasters im ersten Suchschritt entspricht und wobei zur Interpolation mehr Nachbarbildpunkte herangezogen werden als bei einer bilinearen Interpolation,
- in einem dritten Suchschritt wird ausgehend von dem auf Subpelgenauigkeit bestimmten Bewegungsvektor durch eine weitere Interpolationsfilterung mittels des digitalen Filters ein weiter verbesserter Bewegungsvektor ermittelt,

30

35

wobei die Auflösung gegenüber dem zweiten Suchschritt nochmals erhöht wird und die Interpolation basierend auf dem Bildpunktraster mit der Auflösung im zweiten Suchschritt vorgenommen wird.

5

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß zur Interpolationsfilterung beim zweiten Suchschritt ein FIR-Filter verwendet wird mit den Filterkoeffizienten  $CO1 = 161/256$ ,  $CO2 = -43/256$ ,  $CO3 = 23/256$ ,  $CO4 = -8/256$ .

10

3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß zur Interpolationsfilterung beim dritten Suchschritt ein FIR-Filter verwendet wird, mit den FIR-Filterkoeffizienten  $CO1' = 1/2$ ,  $CO2' = 0$ ,  $CO3' = 0$ ,  $CO4' = 0$ .

15

4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß für die Prädiktion von Video-Objekten (VO) die Filterkoeffizienten des digitalen Filters/FIR-Filters für jedes Video-Objekt getrennt aufbereitet werden und in einen Übertragungsbitstrom zu Beginn der Übertragung des jeweiligen Objektes eingefügt werden.

20

25

5. Verfahren nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß zur Codierung eines Bewegungsvektors, insbesondere für eine Übertragung der Wertebereich der zu codierende Bewegungsvektor-Differenzen an eine erhöhte Auflösung angepaßt wird.

15.07.97 Sk/Os

5

ROBERT BOSCH GMBH, 70442 Stuttgart

10

Verfahren zur Erzeugung eines verbesserten Bildsignals bei der Bewegungsschätzung von Bildsequenzen, insbesondere eines Prädiktionssignals für Bewegtbilder mit bewegungskompensierender Prädiktion

15

Zusammenfassung

20

Zur Erzeugung eines verbesserten Bildsignals bei der Bewegungsschätzung erfolgt zuerst eine pelgenaue Bestimmung eines Bewegungsvektors und anschließend eine zweistufige Interpolationsfilterung auf Sub-Pelgenauigkeit. Die Interpolationskoeffizienten werden hinsichtlich einer Aliasing-Reduzierung gewählt. Zur Interpolation werden mehr Nachbarbildpunkte herangezogen als bei üblichen Interpolationsverfahren.

25

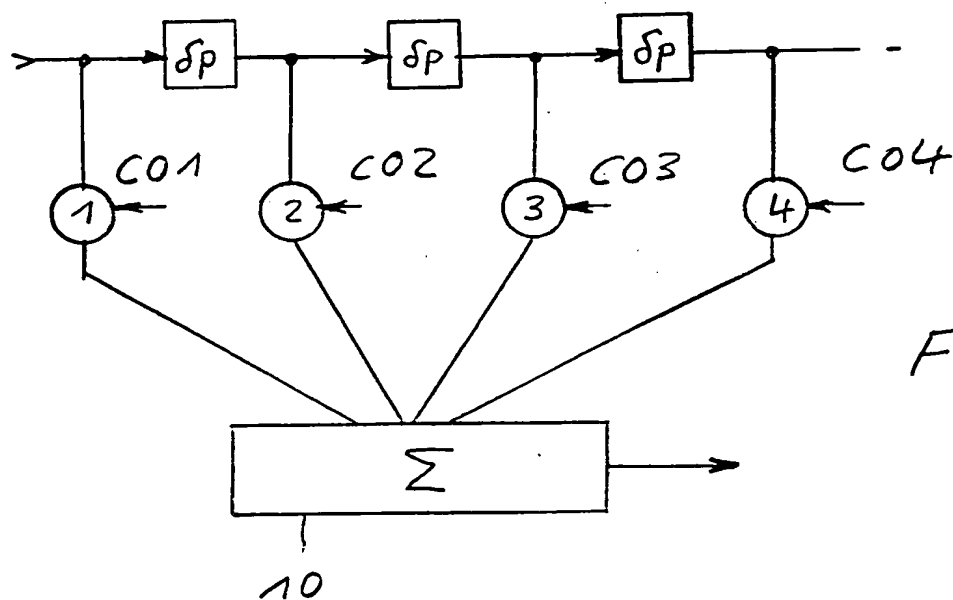
Die Qualität eines Prädiktionssignals für Bewegungsbilder läßt sich verbessern und damit die Codierungseffizienz steigern.

$A \oplus$     $b \bigcirc$     $B +$   
 $c \bigcirc$     $d \bigcirc$   
 $C +$     $D +$

*Fig. 1*

$D_{-4} +$	$D_{-3} +$	$D_{-2} +$	$D_{-1} +$	$D_{+1} +$	$D_{+2} +$	$D_{+3} +$	$D_{+4} +$
$C_{-4} +$	$C_{-3} +$	$C_{-2} +$	$C_{-1} +$	$C_{+1} +$	$C_{+2} +$	$C_{+3} +$	$C_{+4} +$
$B_{-4} +$	$B_{-3} +$	$B_{-2} +$	$B_{-1} +$	$B_{+1} +$	$B_{+2} +$	$B_{+3} +$	$B_{+4} +$
$A_{-4} +$	$A_{-3} +$	$A_{-2} +$	$A_{-1} +$	$b_0 A_{+1} +$	$A_{+2} +$	$A_{+3} +$	$A_{+4} +$
$c_{-4} +$	$c_{-3} +$	$c_{-2} +$	$c_{-1} +$	$d_0 c_{+1} +$	$c_{+2} +$	$c_{+3} +$	$c_{+4} +$
$E_{-4} +$	$E_{-3} +$	$E_{-2} +$	$E_{-1} +$	$E_{+1} +$	$E_{+2} +$	$E_{+3} +$	$E_{+4} +$
$F_{-4} +$	$F_{-3} +$	$F_{-2} +$	$F_{-1} +$	$F_{+1} +$	$F_{+2} +$	$F_{+3} +$	$F_{+4} +$
$G_{-4} +$	$G_{-3} +$	$G_{-2} +$	$G_{-1} +$	$G_{+1} +$	$G_{+2} +$	$G_{+3} +$	$G_{+4} +$
$H_{-4} +$	$H_{-3} +$	$H_{-2} +$	$H_{-1} +$	$H_{+1} +$	$H_{+2} +$	$H_{+3} +$	$H_{+4} +$

*Fig. 2*



-	-	-	-	-	-	-
-	o	-	o	-	o	-
-	-	-	-	-	-	-
-	o	-	x	-	o	-
+	+	+	-	-	-	-
+	o	+	o	-	o	-
+	+	+	-	-	-	-

Fig. 4